Development of a Rail Station Choice Model for NJ TRANSIT

CHERYL ROSEN KASTRENAKES

NJ TRANSIT sought to develop a rail station choice model that would allow the agency to forecast rail ridership to specific stations. This ability would give NJ TRANSIT the opportunity to define its markets as narrowly as possible so that it could best respond to the escalating travel demand occurring in the region. The end result of running the station choice model would be a new boarding-point distribution for each town that would result from either changes in policy variables affecting a given station or from the addition or deletion of stations in a town's "choice set" of stations. A multinomial logit model provided the appropriate framework for analyzing and predicting the station choices of rail users. A data base was created that included each municipality boarding-point pair found in the 1983 ridership origin-destination survey conducted by NJ TRANSIT, the share of riders from each town to each station, and some 20 service and demographic variables. The results of the analysis indicate that station choice is most influenced by the presence of a station in the passenger's town. Access time to the station exerts the next largest influence, followed closely by the frequency of service during the peak hour. Peak travel time and fare on the commuter railroad combined into a generalized cost variable were also significant in determining station choice, but had the smallest effect.

NJ TRANSIT, operator of all commuter rail service and substantial local and commuter express bus services in New Jersey, has been confronted with rapidly escalating travel demand. This increase is caused by sustained and continuing population growth in New Jersey coupled with job growth in Manhattan, a major destination for New Jersey office workers. The mounting pressure this growth has caused on area roads, rail lines, and bus routes suggests the need for careful planning to extend and expand transit services around the state as a means of capturing new travelers, reducing traffic congestion, and supporting the state's economic expansion and vitality.

In response to this challenge, NJ TRANSIT, in partnership with the Port Authority of New York and New Jersey and the New Jersey Department of Transportation, has undertaken a number of studies aimed at identifying the costs and benefits associated with a variety of transit improvements and park-ride expansions (*1*). The key to this work is the analysis of the probable market response to these improvements; that is, shifts of travelers from automobile to transit and among transit modes in relation to the cost of the improvements and how easy they are to implement. Given the large expense associated with even small improvements in commuter rail services, proposals must be critically analyzed to determine whether riders would enjoy a net gain and how they might redistribute themselves among modes, rail lines, and boarding stations in response to new services. NJ TRANSIT needed tools to help predict these responses for the following three types of rail service improvements:

1. New services on existing lines,
2. Extensions of existing lines or construction of new lines in areas not now receiving commuter rail service, and
3. The creation of new stations on overcrowded lines.

Although intuitive analysis of current travel behavior may help predict the impact of such improvements, NJ TRANSIT determined that, given the scope of the projects under consideration, the volatility of New Jersey population growth and travel markets, and the need for a highly structured and objective evaluation process, a statistically valid set of predictive models would provide the only reasonable means of generating traveler response to system changes and subsequent cost performance of project options. Only with such models could reliable estimates of the comparative market value of different projects and their overall cost-effectiveness be understood.

To provide for this market analysis, a mode-shift model was developed that projects traveler response to changes in travel cost, time, or transfers among modes competing in defined corridors. This model helps planners understand the likely mode-shift consequences of a particular rail or bus service improvement.

To drive the mode-split model, however, planners must determine how travelers originating in an area would distribute themselves among boarding stations under the service option to be tested. To mechanically determine this boarding-point distribution for rail system service changes, NJ TRANSIT decided to develop a "station choice" model. This model would incorporate, as appropriate, those factors that determine rail station choice and would provide a means of redistributing travelers, given the proposal for new services available in specified corridors. For example, if travel time from a particular station is reduced, then the number of people using the station will most likely increase. Likewise, if greater frequency is provided from one station than from another, more people will use that station, all else being equal.

Described in this paper is the development of the station choice model, and some insight is given into its value in designing new services that will strengthen transit in New Jersey in support of ongoing growth and development.

New Jersey Transit Corporation, Inc., Department of Planning, McCarter Highway and Market St., P.O. Box 10009, Newark, N.J. 07101.

BACKGROUND

NJ TRANSIT's overall modeling project has entailed determining (a) what factors relate to travelers' choice of mode, and (b) when travelers choose rail, how they choose among boarding stations. With this information, NJ TRANSIT has developed predictive models with which to forecast the likely redistribution of travelers, first among modes and second among boarding stations for rail riders. In many cases, the factors affecting choice among modes and among stations have proven similar; however, the macro- versus micro-scale focus of a corridor-level mode-shift model and a local station choice model required that these factors be accorded different weights in the two applications and that some factors be included in one model that are not material to the other.

The end result of running the station choice model is a new boarding-point distribution for each town supplying riders to the rail system (see Figure 1). This distribution is used as both an input to the mode split model and to apportion the mode-split model's output to generate final boarding-point loadings. The station choice model is run to determine the likely percentage distribution of a body of riders generated at the municipality level among those stations that make up the likely set of boarding-point options for those riders. Presumably, as conditions among competing boarding points change, relative appeal changes as well, and riders "vote with their feet" for the better boarding point. The station choice model predicts only the share of riders from an origin point that will use each station in the choice set, not the absolute number of riders that can be handled.

The new shares calculated from the station choice model are used to change the weights of critical variables in the mode choice model—travel time, number of transfers required en route, and travel cost. After the mode choice model has produced the new ridership for each mode that results from these changes, the station choice model input can be used at the output side of the calculation to apportion these travelers back to the boarding points as actual station loadings (see Figure 2). This information in turn can be translated into parking demands, park-ride requirements, platform and corridor pedestrian flows, and, finally, capital (and perhaps operating) costs of the service improvement.

With the ability to forecast ridership at the station level, NJ TRANSIT will be able to fine tune its park-ride program, station rehabilitation program, and operating schedules. Without a station choice model, the decision of how to weight the variables in the mode choice model when testing changes to the system would be left to planning judgment. The model will drastically reduce the time that would otherwise be needed to determine rider distribution and offer a consistent statistical basis to the process. Together, the two models are critical to NJ TRANSIT's ability to objectively appraise the market and cost potential of key transit system proposals: the Kearny Connection, Secaucus Transfer/Connection, West Shore Railroad, Monmouth/Middlesex/Ocean rail and bus options and others (see Figure 3).

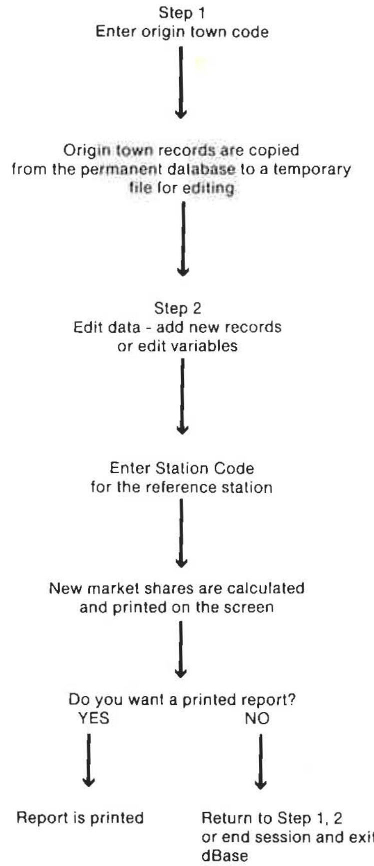

FIGURE 1 Station choice model flowchart.

QUANTITATIVE THEORY

A multinomial logit model provided the appropriate framework for analyzing and predicting the station choices of rail users (2). The standard form of the multinomial logit model is

$$S_i = e^{U_i} \Big/ \sum_{j=1}^{J} e^{U_j}$$

where

- S_i = the share of rail users from a given minor civil division (MCD) who board at station i;
- U_i = the utility associated with the use of station i by rail users from that MCD. Utility is expressed as a linear function of level-of-service variables; and
- J = the number of stations used by rail riders from that MCD.

Because of the availability of regression software for the microcomputer, the logit model was transformed into a form that was linear in parameters and was developed as a regression equation.

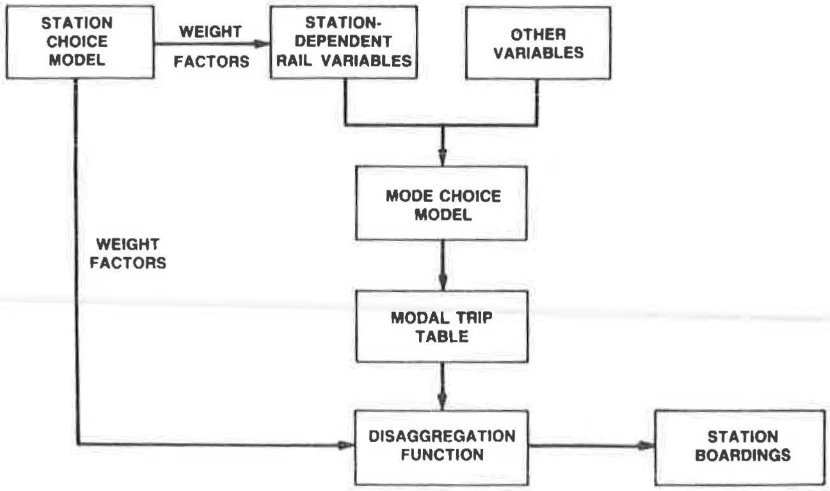

FIGURE 2 Station choice/mode choice flowchart.

VARIABLES AND RESULTS

In 1983, NJ TRANSIT conducted an origin and destination survey of riders on each of its nine commuter rail lines. More than 45,000 survey questionnaires were distributed to 67,000 rail passengers. A total of 26,000 usable surveys were returned, yielding a 58 percent response rate (39 percent of passengers) (*3*). The origin–boarding-point information for this work was obtained from the following questions contained in the survey:

1. Where did you travel from just prior to boarding this train?
2. Is this your residence? Yes or No—I live at _____; and
3. At what station did you board this train?

A data base was created that included each of the municipality–boarding-point pairs found in the 1983 rail survey, the share of riders from each town to each station, and some 20 service and demographic variables.

The variables tested for value in the regression equation included: average access distance, average access time, fare, line-haul distance, density, median household income, transfers, parking availability, parking fees, peak frequency, peak travel time, station location, speed, and cost.

After running regressions with different variable combinations, many of these variables were found to be noncontrolling of rail riders' station choice. The variables that were found to influence station choice were: whether the station was "residential" (i.e., local to the residents selecting it), the access time required to reach the station from the origin municipality, the frequency of rail service from the boarding point, and a generalized cost of the trip from that point. Including these variables as appropriate to their relative strength in driving station choice behavior, the final equation for station choice decisions is

$$\log (S_s/S_r) = 1.5\ (\text{res}_s - \text{res}_r) - .027(t_s - t_r) + .383(f_s - f_r) - .005(C_s - C_r)$$

where S_s is the share of rail users from a given MCD who board at the subject station, and S_r is the share of rail users from a given MCD who board at the reference station. The reference station is always the station that receives the smallest share of riders from a given MCD.

The variables driving station choice derived from the regression, in order of importance, are

res = A dummy variable indicating whether a given station is located in a given MCD and acts as a local station in terms of its proximity to the majority of boarders. 1 = Yes, 0 = No;
t = The access time from the origin MCD to the station. Access time was determined by measuring the shortest route from the center of the residential section of the town to the rail station and converting the distance into time;
f = The frequency of service, trains per hour during the morning peak period; and
c = A generalized cost variable calculated as

Fare + (peak travel time/60 × $8.01)

where $8.01 is one-half of the average hourly wage rate. The empirical work for the model is summarized in Table 1.

The results of the regression analysis indicate that station choice is most influenced by the presence of a station in the passenger's town. The lack of correlation between the local station variable and access time highlights the importance of nonquantitative factors in station choice. This variable seems to capture such intangibles as greater awareness of the services available in the passenger's town and perceived security and knowledge of available parking sites. Access time to the station exerts the next largest influence, followed closely by the frequency of service during the peak hour. The generalized cost variable is also significant in determining station choice but has the smallest effect. Peak travel time and fare were highly correlated and could not be used separately in the equation. The variables were also stronger combined in the cost variable than either was independently.

APPLICATIONS

The following material gives some evidence of the station choice model's use to NJ TRANSIT. NJ TRANSIT is, as part

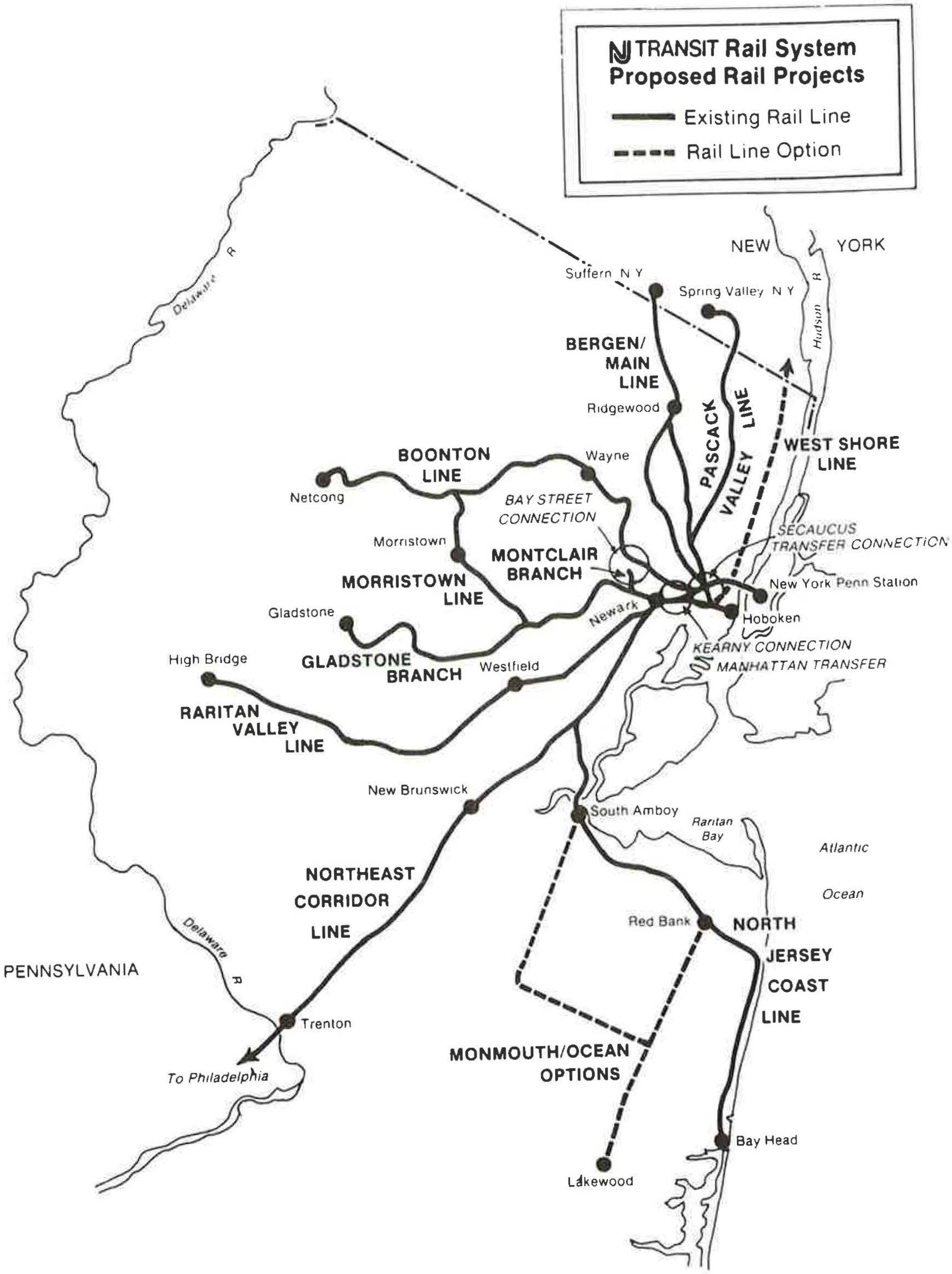

FIGURE 3 Proposed rail projects.

of its overall planning agenda, contemplating the creation of new commuter rail services on branch freight lines that meet current passenger rail lines now in service. One such option in Middlesex County would place new commuter rail service to Newark and Manhattan in an area now served only by buses subject to substantial traffic congestion. Positioned roughly midway between two existing commuter rail services, it is anticipated that the new service will both open up new rail markets and siphon some riders away from each of the parallel routes.

The station choice model was used to test this latter effect as a means of preparing input for the mode choice model and for understanding how travelers might be redistributed if the new service were introduced.

TABLE 1 STATION-CHOICE EQUATION

	Res	Access Time	Frequency	Cost	R^2	F
Parameter estimate	1.5	−.027	.383	−.005	.325	86.6
Standardized estimate	.8	−.493	.434	−.165		
Probability T	.0001	.0001	.0001	.0175		

The results are given in Table 2. The place tested was East Brunswick, a heavily populated town in the center of a major commuter area.

The second column in the table lists the stations that are now used now by East Brunswick commuters and those likely to be used with the introduction of the new rail service. As can be seen by a review of the table, the majority of rail riders originating in East Brunswick (51 percent) use the New Brunswick station to board existing rail services, with lesser proportions using Metropark (16 percent), Metuchen (12 percent), Jersey Avenue (12 percent), and South Amboy (8 percent)—stations located further away. Note that the New Brunswick station is not a resident station but offers the shortest access time from East Brunswick of any of its competitors.

The column Market Share indicates these percentages, with zero as the percentage of East Brunswick rail riders boarding at the three stations that do not currently exist. These other stations are included because they will become available only with the introduction of new rail service and do not exist now.

The last column in the printout, Projected Share, gives the redistribution of East Brunswick rail riders likely to occur with the introduction of new rail service. As can be seen, the new service would not exert a major influence on East Brunswick and would do little to relieve New Brunswick. Because none of the new stations would be local for East Brunswick residents, and none of the existing stations are local, only access time, travel time, and service frequency exert any influence. With New Brunswick still the station with the shortest access time from East Brunswick, it still draws the majority of East Brunswick riders: 42 percent, with the three new stations—Route 18, Cheesequake, and South Brunswick skimming only a handful of the riders from each of the existing stations.

Equivalent tests for each of the municipalities likely to be affected by the new rail service gave widely varying results, depending on the influence of town proximity to the station—the res or local station variable, access time, and the other variables noted previously. The combined weighted average of these township station assignments form the input by which the relative appeal of the new rail service in drawing riders from bus and automobile can be determined by the mode choice model. The station choice model outputs listed in Table 2 are then used to reassign this predicted mode shift back to the station boarding point to guide park-ride analysis and final project costing.

CONCLUSION

The aim of this work has been to determine which factors most influence station choice and to develop a model from these results. The variables with the strongest influence were the presence of a station in the rider's town, access time, peak frequency and travel time, and fare combined into a generalized cost variable.

Unfortunately, some variables that may have proved important were difficult to work with and had to be dropped from the analysis. Parking fee and parking availability are two examples. Parking fee was a significant variable but the coefficient had a positive sign. This implied that as parking fees increase at a station, so would the desire to choose that station. This result is counterintuitive and would distort the model's predictive capabilities. The higher parking fees do not create the demand. In reality, the higher fees are created as a result of a combination of factors that include: (a) high demand for the station because of good service or easy access, or both; (b) a scarcity of parking alternatives or available land to develop parking on the rail system; (c) a heavily congested, and therefore inconvenient, highway system throughout much of the state as NJ TRANSIT's alternative; and (d) high household incomes among the majority of NJ TRANSIT's rail passengers. The same type of situation occurred with the parking-availability variable, where lack of parking availability is because many travelers board there rather than that the lack of parking results in travelers choosing the station.

The results imply that NJ TRANSIT needs to be aware of and responsive to people's need for convenience in their desire both to be near a station and to have frequent service. These results confirm most transportation studies in that a unit of access time and wait time are considered more important to

TABLE 2 STATION CHOICE REPORT

NJ Transit Code	Town	Station Code	Station	Market Share	
				Actual	Projected
23020	East Brunswick	17	South Amboy	8	6
23020	East Brunswick	903	Jersey Ave.	12	10
23020	East Brunswick	906	Metuchen	12	10
23020	East Brunswick	907	Metropark	16	13
23020	East Brunswick	904	New Brunswick	51	42
23020	East Brunswick	981	Route 18	0	7
23020	East Brunswick	980	Cheesequake	0	6
23020	East Brunswick	913	South Brunswick	0	5

riders than a unit of in-vehicle line-haul time. It should not be inferred, however, that only local rail stations can be successful. It has been proven on the NJ TRANSIT rail system that stations with easy highway access, high quality service, and abundant parking can be successful regional rail stations.

ACKNOWLEDGMENTS

Appreciation is extended to Kevin Neels of Charles River Associates and Steve Jurow of NJ TRANSIT for their assistance on this project; Neels, for his invaluable technical knowledge which helped to get the project started, and Jurow for his guidance and support.

REFERENCES

1. *Meeting New Jersey's Growth Challenge—A Transit Plan for the 1990's.* NJ TRANSIT, Planning Department, Newark, N.J., 1987.
2. *Modeling Trans-Hudson Travel Demand.* Charles River Associates Incorporated Report No. 101.03, Boston, Mass., 1987, pp. 95–103.
3. *NJ TRANSIT Rail Passenger Survey—Summary Report.* NJ TRANSIT, Planning Department. Newark, N.J., 1984.

Publication of this paper sponsored by Committee on Rail Transit Systems.